

An instanton approach to large N Harish-Chandra-Itzykson-Zuber integrals

J. Bun^{1,2,3}, J. P. Bouchaud¹, S. N. Majumdar², M. Potters¹

¹ *Capital Fund Management, 23–25, rue de l'Université, 75007 Paris*

² *CNRS, LPTMS, Bat. 100, Université d'Orsay, 91405 Orsay Cedex and*

³ *DeVinci Finance Lab, Pôle Universitaire Léonard de Vinci, 92916 Paris La Défense, France.*

We reconsider the large N asymptotics of Harish-Chandra-Itzykson-Zuber integrals. We provide, using Dyson's Brownian motion and the method of instantons, an alternative, transparent derivation of the Matytsin formalism for the unitary case. Our method is easily generalised to the orthogonal and symplectic ensembles. We obtain an explicit solution of Matytsin's equations in the case of Wigner matrices, as well as a general expansion method in the dilute limit, when the spectrum of eigenvalues spreads over very wide regions.

The ability to perform explicit calculations of sums and integrals is at the heart of many groundbreaking progress in theoretical physics, in particular field theory or statistical mechanics. In that respect, the so-called Harish-Chandra-Itzykson-Zuber (HCIZ) integral [1, 2] is among the most beautiful results, and has found several applications in many different fields, including Random Matrix Theory, disordered systems or quantum gravity (for a particularly insightful introduction, see [3]). The generalized HCIZ integral $\mathcal{I}_\beta(A, B)$ is defined as:

$$\mathcal{I}_\beta(A, B) = \int_{G(N)} \mathcal{D}\Omega e^{\frac{\beta N}{2} \text{Tr} A \Omega B \Omega^\dagger}, \quad (1)$$

where the integral is over the (flat) Haar measure of the compact group $\Omega \in G(N) = O(N), U(N)$ or $Sp(N)$ in N dimensions and A, B are arbitrary $N \times N$ symmetric (resp. hermitian or symplectic) matrices. The parameter β is the usual Dyson “inverse temperature”, with $\beta = 1, 2$ or 4 respectively for the three groups. In the unitary case $G(N) = U(N)$ and $\beta = 2$, it turns out that the HCIZ integral can be expressed exactly, for all N , as the ratio of determinants that depend on A, B and additional N -dependent prefactors:

$$\mathcal{I}_{\beta=2}(A, B) = \frac{c_N}{N^{(N^2-N)/2}} \frac{\det((e^{N\nu_i\lambda_j})_{1 \leq i, j \leq N})}{\Delta^2(A)\Delta^2(B)} \quad (2)$$

with $\{\nu_i\}$, $\{\lambda_i\}$ the eigenvalues of A and B , $\Delta(A) = \prod_{i < j} |\nu_i - \nu_j|$ the Vandermonde determinant of A (and similarly for $\Delta(B)$); and $c_N = \prod_{i=1}^N i!$.

The situation is however somewhat paradoxical. Although the HCIZ result is fully explicit for $\beta = 2$, the expression in terms of determinants is highly non trivial and quite tricky. For example, the expression becomes degenerate (0/0) whenever two eigenvalues of A (or B) coincide. Also, as is well known, determinants contain $N!$ terms of alternating signs, which makes their order of magnitude very hard to estimate *a priori*. This difficulty appears clearly when one is interested in the large N asymptotics of HCIZ integrals, for which one would naively expect to have a simplified, explicit expression as a functional $F_2(\rho_A, \rho_B) = \lim_{N \rightarrow \infty} N^{-2} \ln \mathcal{I}_{\beta=2}(A, B)$ of

the eigenvalue densities $\rho_{A,B}$ of A, B . But even this large N limit turns out to be highly non trivial. In a remarkable paper, Matytsin [4] suggested a mapping to a non-linear hydrodynamical problem in one-dimension, the solution of which gives in principle access to $F_2(\rho_A, \rho_B)$. Matytsin's result for $N \rightarrow \infty$ was later shown by Guionnet & Zeitouni [5] to be mathematically rigorous. Still, neither Matytsin's, nor Guionnet & Zeitouni's derivation is very transparent (at least to our eyes); furthermore, the generalisations of Matytsin's method to the orthogonal case – which appears in many interesting statistical mechanics problems – or the symplectic case has not been worked out. In this letter, we recover Matytsin's equations using a rather straightforward instanton approach to the Dyson Brownian motion that describes the (fictitious) dynamics of eigenvalues connecting ρ_A to ρ_B . Our approach is easily adapted to *arbitrary values* of β , including the orthogonal case which yields Zuber's “ $\frac{1}{2}$ -rule” when $N \rightarrow \infty$, i.e. $F_1(\rho_A, \rho_B) = F_2(\rho_A, \rho_B)/2$ [6]. We then solve exactly Matytsin's equation in two particular cases (i) both ρ_A and ρ_B are Wigner semi-circle distributions (of arbitrary widths $\sigma_{A,B}$); (ii) ρ_A and ρ_B are arbitrary, but with diverging widths $\sigma_{A,B} \rightarrow \infty$. We compare our results with the small- σ expansion obtained in [7].

Our main idea is to study, using the method of instantons, the large deviations of the Dyson Brownian motion of eigenvalues that brings an initial distribution of eigenvalues ρ_A to a final distribution ρ_B (see Fig. 1). This occurs with a probability that is exponentially small, $\propto \exp(-N^2 S)$, with a rate S that we are able to relate directly to the HCIZ integral – see below. Suppose that one adds to a certain matrix A small random Gaussian Hermitian matrices of variance dt/N . It is well known that in the $dt \rightarrow 0$ limit, the eigenvalues $\{x_i\}$ of the time-dependent matrix evolve according to (see [8]):

$$dx_i = \sqrt{\frac{2}{\beta N}} dW + \frac{1}{N} dt \sum_{j \neq i} \frac{1}{x_i - x_j}, \quad (3)$$

where $W(t)$ is the standard Brownian motion and we set henceforth $\beta \equiv 2$, corresponding to unitary matrices. The calculation of S can be done using two different (but

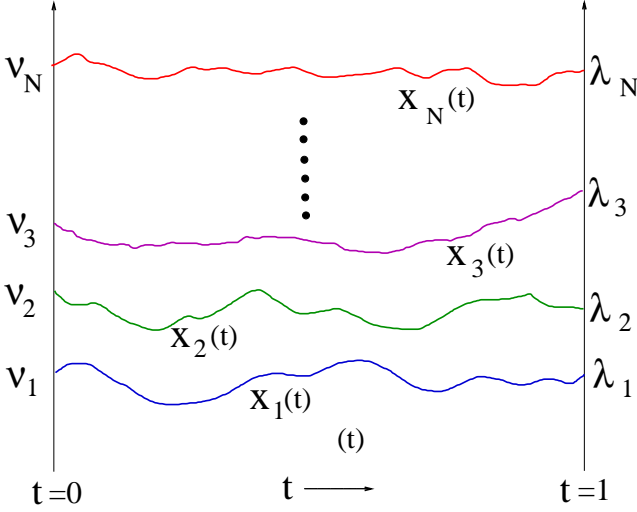


Figure 1: Dyson Brownian motion transporting the initial distribution ρ_A of the eigenvalues of A to the final distribution ρ_B of the eigenvalues of B , in a (fictitious) time $t = 1$.

complementary) languages: that of particle trajectories and that of densities, using the Dean-Kawasaki formalism. We start with the particle point of view, and sketch the density functional method later. We introduce the total potential energy $U(\{x_i\}) = -\frac{1}{N} \sum_{i < j} \ln |x_i - x_j|$, and the corresponding “force” $f_i = -\partial_{x_i} U$. The probability of a given trajectory for the N Brownian motions between time $t = 0$ and time $t = 1$ is given by (see Fig. 1): [17]

$$\begin{aligned} \mathcal{P}(\{x_i(t)\}) &= \mathcal{N} \exp - \left[\frac{N}{2} \int_0^1 dt \sum_i (\dot{x}_i + \partial_{x_i} U)^2 \right] \\ &\equiv \mathcal{N} e^{-N^2 S}. \end{aligned} \quad (4)$$

The action $S = S_1 + S_2$ contains a total derivative equal, in the continuum limit, to:

$$S_1 = -\frac{1}{2} \left[\int dx dy \rho_Z(x) \rho_Z(y) \ln |x - y| \right]_{Z=A}^{Z=B} \quad (5)$$

and:

$$S_2 = \frac{1}{2N} \int_0^1 dt \sum_{i=1}^N [\dot{x}_i^2 + (\partial_{x_i} U)^2] \quad (6)$$

The “instanton” trajectory that dominates the probability for large N is such that the functional derivative with respect to all $x_i(t)$ is zero (see e.g. [9]):

$$-2 \frac{d^2 x_i}{dt^2} + 2 \sum_{\ell=1}^N \partial_{x_i, x_\ell}^2 U \partial_{x_\ell} U = 0 \quad (7)$$

which leads, after a few algebraic manipulations, to:

$$\frac{d^2 x_i}{dt^2} = -\frac{2}{N^2} \sum_{\ell \neq i} \frac{1}{(x_i - x_\ell)^3}. \quad (8)$$

This can be interpreted as the motion of unit mass particles, accelerated by an *attractive* force that derives from an effective two-body potential $\phi(r) = -(Nr)^{-2}$. The hydrodynamical description of such a fluid is given by the Euler equations for the density $\rho(x, t)$ and the velocity field $v(x, t)$: [18]

$$\partial_t \rho(x, t) + \partial_x [\rho(x, t) v(x, t)] = 0, \quad (9)$$

and

$$\partial_t v(x, t) + v(x, t) \partial_x v(x, t) = -\frac{1}{\rho(x, t)} \partial_x P(x, t), \quad (10)$$

where $P(x, t)$ is the pressure field which reads, from the virial formula in one dimension [10, p. 138]:

$$P = \rho T - \frac{1}{2} \rho \sum_{\ell \neq i} |x_i - x_\ell| \phi'(|x_i - x_\ell|) \approx -\frac{\rho}{N^2} \sum_{\ell \neq i} \frac{1}{(x_i - x_\ell)^2}, \quad (11)$$

because the fluid is at an effective temperature $T = 1/N$ (see below). Now, using the same argument as Matytsin [4], i.e., writing $x_i - x_\ell \approx (i - \ell)/(N\rho)$ and $\sum_{n=1}^\infty n^{-2} = \frac{\pi^2}{6}$, one finally finds: [19]

$$P(x, t) = -\frac{\pi^2}{3} \rho(x, t)^3, \quad (12)$$

and therefore Matytsin’s equations for ρ and v . Plugging this back in the action S , and going to the continuous limit, one also finds:

$$S_2 \approx \frac{1}{2} \int dx \rho(x, t) \left[v^2(x, t) + \frac{\pi^2}{3} \rho^2(x, t) \right] \quad (13)$$

which is exactly Matytsin’s action [4]. Finally, the probability $\mathcal{P}(\{\lambda_i\}|\{\nu_i\})$ to observe the set of eigenvalues $\{\lambda_i\}$ of B for a given set of eigenvalues ν_i for A is proportional to $\exp[-N^2(S_1 + S_2)]$ where S_2 is obtained by plugging into Eq. (13) the solution of the Euler equations (9, 10), with $v(x, t = 0)$ chosen in such a way that $\rho(\nu, t = 0) = \rho_A(\nu)$ and $\rho(\lambda, t = 1) = \rho_B(\lambda)$.

Now, the idea is to interpret the HCIZ integrand in the unitary case, $\exp[N \text{Tr} A U B U^\dagger]$, as a part of the propagator of the diffusion operator in the space of Hermitian matrices. Indeed, adding to A a small random Gaussian Hermitian matrices of variance dt/N , the probability to end up with matrix B in a time $t = 1$ is $\mathcal{P}(B|A) \propto \exp[-N/2 \text{Tr}(A - B)^2]$. Writing $B = V \Lambda V^\dagger$ with $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_N)$, the change of variables, as is well known, induces a probability measure on $\{\lambda_i\}$ alone that includes a Vandermonde determinant $\Delta^2(B) = \prod_{i < j} |\lambda_i - \lambda_j|^2$. Since the conditional distribution of $\{\lambda_i\}$ is obviously invariant under $B \rightarrow U B U^\dagger$ where U is an arbitrary unitary transformation, we get

another expression for $\mathcal{P}(\{\lambda_i\}|\{\nu_i\})$: [20]

$$\begin{aligned} & \mathcal{P}(\{\lambda_i\}|\{\nu_i\}) \\ & \propto \prod_{i < j} |\lambda_i - \lambda_j|^2 \int \mathcal{D}U \exp \left[-\frac{N}{2} \text{Tr}(A - UBU^\dagger)^2 \right] \\ & \propto \Delta^2(B) \exp \left[-\frac{N}{2} (\text{Tr} A^2 + \text{Tr} B^2) \right] \mathcal{I}_2(A, B). \end{aligned} \quad (14)$$

Comparing this last expression for $\beta = 2$ with the above calculation, and taking care of the proportionality coefficients, we get as a final expression for $F_{\beta=2}(A, B) = \lim_{N \rightarrow \infty} N^{-2} \ln \mathcal{I}_2(A, B)$:

$$\begin{aligned} F_2(A, B) = & -\frac{3}{4} - S_2(A, B) + \frac{1}{2} \int dx x^2 (\rho_A(x) + \rho_B(x)) \\ & - \frac{1}{2} \int dx dy [\rho_A(x)\rho_A(y) + \rho_B(x)\rho_B(y)] \ln |x - y|, \end{aligned} \quad (15)$$

which is, apart from the $-3/4$ term which comes from the prefactor in Eq. (2), precisely Matytsin's result [4]. Now, the whole calculation above can be repeated for the $\beta = 1$ (orthogonal group) or $\beta = 4$ (symplectic group) with the final (simple) result $F_\beta(A, B) = \beta F_2(A, B)/2$. This coincides with the result obtained by Zuber in the orthogonal case $\beta = 1$ [6] (see also [11]).

We now briefly explain how to obtain the same result using the Dean-Kawasaki framework [12, 13]. As shown by Dean [13], the density $\rho(x, t)$ of interacting particles obeying the Langevin equation (3) is found to satisfy the (functional) Langevin equation $\partial_t \rho(x, t) + \partial_x J(x, t) = 0$, with:

$$\begin{aligned} J(x, t) = & \frac{1}{N} \xi(x, t) \sqrt{\rho(x, t)} - \frac{1}{2N} \partial_x \rho(x, t) \\ & - \rho(x, t) \int dy \partial_x V(x - y) \rho(y, t), \end{aligned} \quad (16)$$

where $V(r) = -\ln r$ is the two-body interaction potential, $\xi(x, t)$ is a normalized Gaussian white noise (in time and in space) and unlike in [13], we define $\rho(x, t) = \frac{1}{N} \sum_{i=1}^N \delta(x - x_i(t))$. One can again write the weight of histories of $\{\rho(x, t)\}$ using Martin-Siggia-Rose path integrals. This reads:

$$\mathcal{P}(\{\rho(x, t)\}) \propto \left\langle \int \mathcal{D}\psi e^{\left[\int_0^1 dt \int dx N^2 i\psi(x, t) (\partial_t \rho + \partial_x J) \right]} \right\rangle_\xi \quad (17)$$

Performing the average over ξ gives the following action (and renaming $-i\psi \rightarrow \psi$):

$$\begin{aligned} \mathcal{S} = & N^2 \int_0^1 dt \int dx [\psi \partial_t \rho + f(x, t) \rho \partial_x \psi \\ & - \frac{\psi}{2N} \partial_{xx}^2 \rho + \frac{1}{2} \rho (\partial_x \psi)^2] \end{aligned} \quad (18)$$

with $f(x, t) = \int dy \partial_x V(x - y) \rho(y, t)$. Taking functional derivatives with respect to ρ and ψ then leads to the

following set of equations:

$$\partial_t \rho = \partial_x (\rho f) + \partial_x (\rho \partial_x \psi) + \frac{1}{2N} \partial_{xx}^2 \rho \quad (19)$$

and

$$\begin{aligned} \partial_t \psi - \frac{1}{2} (\partial_x \psi)^2 = & f \partial_x \psi - \frac{1}{2N} \partial_{xx}^2 \psi \\ & - \partial_x \int dy V(x - y) \rho(y, t) \partial_y \psi(y, t). \end{aligned} \quad (20)$$

The Euler-Matystin equations are recovered, after a little work, by setting $v(x, t) = -f(x, t) - \partial_x \psi(x, t)$. One can finally check [14] that the \mathcal{S} coincides with S when using the equation of motion satisfied by ρ, ψ and f . Note that this second method gives rise to additional “diffusion” terms, of order $1/N$, which lead to a viscosity term in the velocity equation. This second method might therefore be more adapted to search for subleading corrections (in N^2) to the action.

Somewhat surprisingly, Matytsin's formalism has not been exploited to find explicit solutions for $F_\beta(A, B)$ in some special cases. One fully solvable case is when A and B have Wigner semi-circle spectra, $\rho_A(\nu) = \sqrt{4\sigma_A^2 - \nu^2}/2\pi\sigma_A^2$, and similarly for ρ_B , with a width σ_B . One can first note that since trivially $F_\beta(A, B) = F_\beta(A/z, zB)$, one can always choose $z = \sqrt{\sigma_A/\sigma_B}$ and set $\sigma_A = \sigma_B = \sigma$. The second remark is that the Euler-Matytsin equations can be solved by choosing $\rho^2(x, t) = \alpha(t) + \gamma(t)x^2$ and $v(x, t) = b(t)x$, which leads to ordinary differential equations for α, γ and b . The final solution is that $\rho(x, t)$ is a Wigner semi-circle for all t , with a width $\Sigma(t)$ given by:

$$\Sigma^2(t) = \sigma^2 + gt(1 - t), \quad g = \sqrt{1 + 4\sigma^4} - 2\sigma^2, \quad (21)$$

and $b(t) = (t - 1/2)g/\Sigma^2(t)$. Note that $\Sigma^2(t = 0) = \Sigma^2(t = 1) = \sigma^2$, as it should be. Injecting these expressions into Eqs. (13, 15) finally leads to (with $\sigma^2 = \sigma_A \sigma_B$):

$$F_{2,W}(A, B) = \frac{1}{2} \left[\sqrt{4\sigma^4 + 1} - 1 - \log \left(\frac{1 + \sqrt{4\sigma^4 + 1}}{2} \right) \right]. \quad (22)$$

For arbitrary matrices A, B , the narrow spectra limit (corresponding to $\sigma \rightarrow 0$) has been worked out by Collins [7]. Specializing his general result to the case of Wigner matrices, one finds:

$$F_{2,W}(A, B) \underset{\sigma \rightarrow 0}{=} \frac{\sigma^4}{2} - \frac{\sigma^8}{4} + \frac{\sigma^{12}}{3} + \mathcal{O}(\sigma^{16}), \quad (23)$$

which coincides with the small σ expansion of Eq. (22). In the opposite limit $\sigma \rightarrow \infty$, we find from Eq. (22):

$$F_{2,W}(A, B) \underset{\sigma \rightarrow \infty}{=} \sigma^2 - \ln(\sigma) - \frac{1}{2} - \frac{\sigma^{-2}}{8} + \frac{\sigma^{-6}}{384} + \mathcal{O}(\sigma^{-10}). \quad (24)$$

This limit can be called the dilute limit and can be studied in full generality, since the solution of the Euler-Matytsin equations can be constructed as a power series of $\varepsilon = 1/\sigma$, where we define $\sigma^2 \equiv \int dx x^2 \rho_A(x)$ (we choose here, without loss of generality, $\text{Tr} A = \text{Tr} B = 0$, and rescale the matrices A, B appropriately such that both have the same variance σ^2). In the case where $\rho_A = \rho_B$ but of arbitrary shape (but provided ρ_A vanishes at the edge of the spectrum), our final result to order ε^6 reads:

$$\begin{aligned} F_2(A, A) &\underset{\varepsilon \rightarrow 0}{=} \int dx x^2 \rho_A(x) \\ &- \int dx dy \rho_A(x) \rho_A(y) \ln |x - y| - \frac{3}{4} - \frac{\pi^2}{6} \int dx \rho_A^3(x) \\ &+ \frac{\pi^4}{24} \int dx \rho_A^3(x) \rho_A'(x)^2 + \mathcal{O}(\varepsilon^{10}) \end{aligned} \quad (25)$$

which is identical to Eq. (24) when ρ_A is a Wigner semi-circle, but holds more generally. Note that terms appear in order of importance in the above formula.

The general expression for $\rho_A \neq \rho_B$ is cumbersome and will be given in a longer version of this work [14]. To order ε^2 , the result reads:

$$\begin{aligned} F_2(A, B) &\underset{\varepsilon \rightarrow 0}{=} \int_0^1 dp X_A(p) X_B(p) \\ &- \frac{1}{2} \int dx dy \rho_A(x) \rho_A(y) \ln |x - y| \\ &- \frac{1}{2} \int dx dy \rho_B(x) \rho_B(y) \ln |x - y| - \frac{3}{4} \\ &- \frac{\pi^2}{6} \int_0^1 dp \rho_A(X_A(p)) \rho_B(X_B(p)) + \mathcal{O}(\varepsilon^6) \end{aligned} \quad (26)$$

where $X_Z(p)$ is such that $p = \int_{-\infty}^X du \rho_Z(u) \in [0, 1]$. For $A = B$, one recovers Eq. (25) by changing variables back from p to x , with the Jacobian $dp/dx = \rho_A(x)$. The leading term in the above expansion is in fact $\int_0^1 dp X_A(p) X_B(p)$ and is easy to interpret: it comes from the fact that in the limit $\sigma \rightarrow \infty$, HCIZ integrals Eq. (1) are dominated by the matrix Ω that diagonalizes B in the diagonal base of A (and the corresponding eigenvalues $\{\lambda\}, \{\nu\}$ are ordered).

The main achievements of this work are two-fold: we first rederived the large N asymptotics of HCIZ integrals, first obtained by Matytsin, using Dyson's Brownian motion and the method of instantons. We also provided an exact, explicit solution for the case of Wigner matrices, as well as a general expansion method in the dilute limit, when the eigenvalue spectra spread over very wide regions. Beyond providing a relatively straightforward

and transparent interpretation of Matytsin's method, our work could provide a valuable starting point to obtain new results, such as the generalisation to other ensembles (Orthogonal, Symplectic, Wishart) but also to understand the structure of subleading (in N^2) corrections. Our explicit results in the dilute limit should also be useful for applications, such as, for example, the Bayesian estimate of large correlation matrices using empirical data [15].

We thank J. Bonart and J. B. Zuber for useful comments and suggestions. S.N.M acknowledges support from ANR grant 2011-BS04-013-01 WALKMAT.

-
- [1] Harish-Chandra, American Journal of Mathematics pp. 87–120 (1957).
 - [2] C. Itzykson and J.-B. Zuber, Journal of Mathematical Physics **21**, 411 (1980).
 - [3] T. Tao, <http://terrytao.wordpress.com/2013/02/08/the-harish-chandra-itzykson-zuber-integral-formula/>.
 - [4] A. Matytsin, Nuclear Physics B **411**, 805 (1994).
 - [5] A. Guionnet and O. Zeitouni, Journal of Functional Analysis **188**, 461 (2002).
 - [6] J.-B. Zuber, Journal of Physics A: Mathematical and Theoretical **41**, 382001 (2008).
 - [7] B. Collins, International Mathematics Research Notices **2003**, 953 (2003).
 - [8] F. J. Dyson, Journal of Mathematical Physics **3**, 1191 (1962).
 - [9] A. Bray and A. McKane, Physical Review Letters **62**, 493 (1989).
 - [10] M. Le Bellac, F. Mortessagne, and G. George Batrouni, *Equilibrium and non-equilibrium statistical thermodynamics* (Cambridge University Press, 2006).
 - [11] B. Collins, A. Guionnet, and E. Maurel-Segala, Advances in Mathematics **222**, 172 (2009).
 - [12] K. Kawasaki, Journal of Physics A: Mathematical and General **6**, 1289 (1973).
 - [13] D. S. Dean, Journal of Physics A: Mathematical and General **29**, L613 (1996).
 - [14] J. Bun, J.-P. Bouchaud, S. N. Majumdar, and M. Potters (forthcoming).
 - [15] J. Bun, J.-P. Bouchaud, and M. Potters (in preparation).
 - [16] K. Johansson, Communication in Mathematical Physics **215**, 683 (2001).
 - [17] We neglect the Jacobian which is small in the large N (small temperature) limit, as usual.
 - [18] Actually, the Dean-Kawasaki formalism allows one to see that a viscosity term, of order N^{-1} , is in fact also present – see below.
 - [19] This also implicitly assumes that assuming that the density ρ vanishes at the edges of the spectrum.
 - [20] A more rigorous derivation in the unitary case $\beta = 2$, that includes all prefactors, uses Johansson's formula [16].